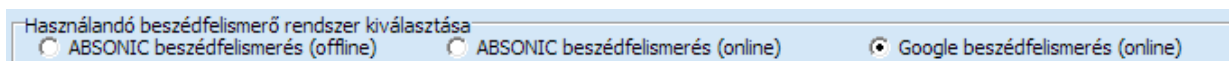


Beszédfelismerés beállításai: Funkciók – Beállítások – Beszédfelismerés beállításai

A program lehetővé teszi egy elkészült hangfelvétel utólagos vagy hangrögzítéssel egy időben történő (kvázi valós idejű) beszédfelismerését és a felismert szöveg fordítását.

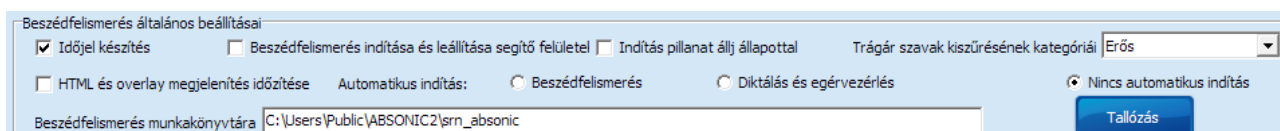
1. Használandó beszédfelismerő rendszer kiválasztása



- **ABSONIC beszédfelismerés (offline)** kiválasztása esetén az ABSONIC saját beszédfelismerő rendszere lesz kiválasztva, amely azon a saját számítógépen fut, amelyen az ABSONIC is.
- **ABSONIC beszédfelismerés (online)** kiválasztása esetén az ABSONIC beszédfelismerő rendszere lesz használva, amely egy másik (szerver) számítógépen fut.
- **Google beszédfelismerés (online)** kiválasztása esetén a Google internetes beszédfelismerő rendszere lesz kiválasztva. Használatához előfizetésre van szükség, mely megrendelhető webáruházunkból vagy a felhasználó maga is csatlakozhat a Google Cloud rendszerhez.

2. Beszédfelismerés általános beállításai

Itt azok a funkciók állíthatók be, amelyek hatása mind a három beszédfelismerő rendszer működését szabályozzák.



2.1. Időjel készítés

- Bekapcsolásakor a beszédfelismerés eredményeként létrejött dokumentum és hangfelvétel is időjelekkel lesz ellátva. Az időjelek segítségével a hangfelvétel és a dokumentum szövegének megfelelő pontjai összekapcsolódnak. Így a hangfelvétel adott pontjára lépve a dokumentum szövegében is a hozzátartozó tartalomhoz ugrik a rendszer és megjeleníti azt.

2.2. Beszédfelismerés indítása és leállítása segítő felületekkel

- Bekapcsolása esetén a beszédfelismerés indításakor **Pillanat állj** állásba kapcsol a rendszer és egy külön gomb megnyomása után indul csak a beszédfelismerés folyamata. A **Pillanat állj** állapotban van lehetőség a beszédfelismerés indulása előtt beállítani a betűtípust, betű méretet, háttér színt, stb.

-
- Bekapcsolása esetén a beszédfelismerés leállításakor egy választó felület jelenik meg, ami segíti a további művelet kiválasztását, például hangfelvétel mentése, dokumentum mentése, felismert szöveg megjelenítése, hangfelvétel lejátszása felismert szöveg megjelenítéssel, stb.

2.3. Indítás pillanat állj állapottal

- Bekapcsolása esetén a beszédfelismerés indításakor **Pillanat állj** állásba kapcsol a rendszer és egy külön gomb megnyomása után indul a beszédfelismerés folyamata. A **Pillanat állj** állapotban van lehetőség a beszédfelismerés indulása előtt beállítani a betűtípust, betű méretet, háttér színt, stb. Viszont a beszédfelismerés leállítása normál módon történik, az előző pontban ismertetett segítő felület nélkül.

2.4. Trágár szavak kiszűrésének kategóriái

- Itt beállíthatjuk a beszédfelismerés folyamán, a trágár szavak kiszűrését végző funkció működésének a kategóriáit. A kiszűrt szavak minden karaktere kicsillagozva jelenik meg.
- A szűrési kategóriák:
 - Erős
 - Közepes
 - Gyenge
 - Kikapcsolva – Minden felismert, trágár szó megjelenik.

2.5. HTML és feliratozás megjelenítés időzítése

- Bekapcsolása esetén a beszédfelismerés (és fordítás) eredményének valós idejű megjelenítésekor a szöveg megjelenési idejét az ABSONIC szoftver automatikusan szabályozza, tehát a hosszabb szövegeket tovább tartja fent, mint a rövideket. Szöveg megjelenítéskor, videó-prezentáció feliratozáskor ajánlott a funkció használata.

2.6. Automatikus indítás

Itt beállíthatjuk, hogy az ABSONIC program indításakor melyik beszédfelismerési szolgáltatás induljon el automatikusan.

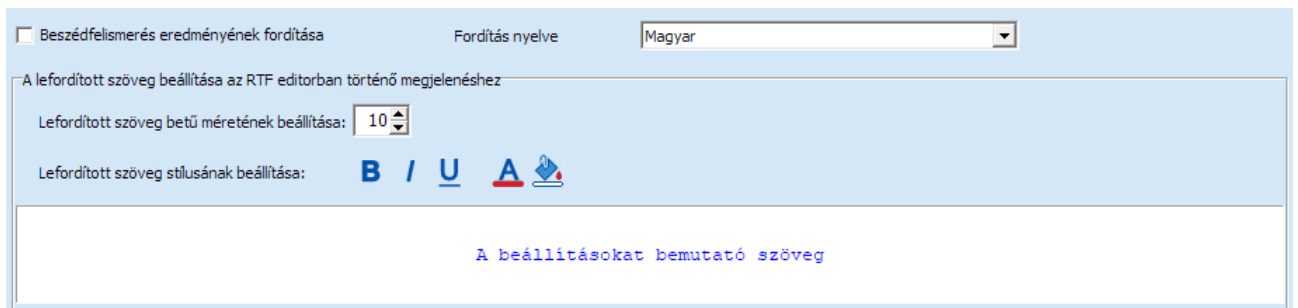
- Beszédfelismerés – A beszédfelismerés indul el automatikusan, a felismerés eredményét a beépített RTF szövegszerkesztőben megjelenítve.
- Diktálás és egérvezérlés – A diktálás és egérvezérlő funkció indul el automatikusan.
- Nincs automatikus indítás – Nem indul automatikusan semmilyen beszédfelismerő funkció






2.7. Beszédfelismerés munkakönyvtára

Itt azt a könyvtárat (mappát) kell kitallózni és beállítani, ahol a beszédfelismerő rendszer állományai találhatóak, illetve az esetleges esemény naplózás történik. Ez alapesetben az \ABSONIC2\srn_absonic mappa.

3. Fordítás

Itt a beszéd felismerés eredményének lefordításához szükséges beállításokat végezhetjük el. A fordítás funkció a **Google beszéd felismerés (online)** beállításainál a Google API kulcs megadását igényli, különben nem működik!



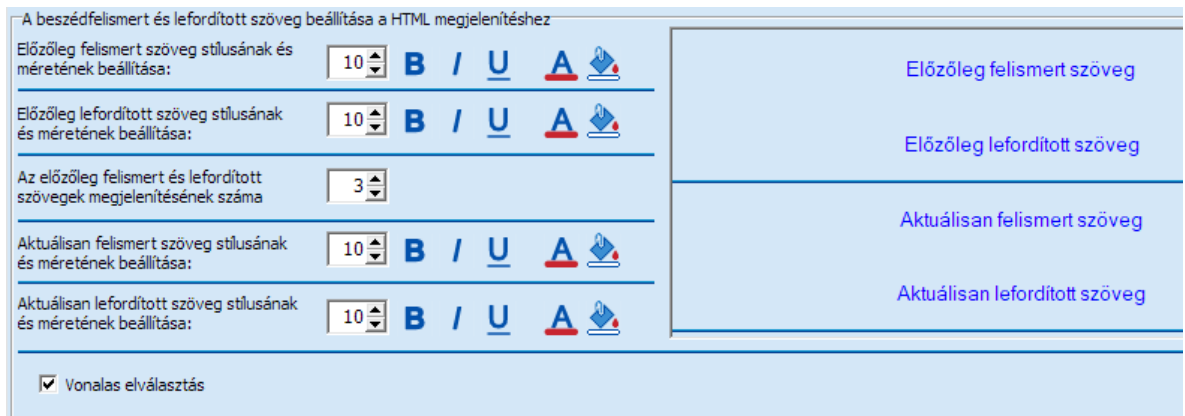
- Be és kikapcsolható a beszéd felismerés eredményének fordítása**
- Fordítás nyelve:** több, mint 100 nyelv közül választhatjuk ki azt a nyelvet, amelyre a beszéd felismerés eredményét fordítani fogja a rendszer.
- Lefordított szöveg betű méretének beállítása:** az RTF editorban (szövegszerkesztőben) megjelenített, lefordított szöveg betűmérete.
- Lefordított szöveg stílusának beállítása**
 -  gomb - félkövér (bold) stílus
 -  gomb - dőlt (italic) stílus
 -  gomb - aláhúzott stílus
 -  gomb - a lefordított szöveg betűszínének beállítása
 -  gomb - a lefordított szöveg hátterszínének beállítása
- A beállításokat egy bemutató szöveg segítségével lehet ellenőrizni

A beállításokat bemutató szöveg







4. HTML megjelenítés

Böngészőben megjeleníthetjük az aktuális és ezt közvetlenül megelőző felismert és lefordított szöveget. A lefordított szövegek csak a fordítás funkció bekapcsolása esetén jelennek meg!

- Egy böngésző címsorába az ABSONIC szoftvert futtató számítógép nevét vagy IP címét írva jeleníthetjük meg a beszéd felismerés és fordítás eredményét.
A böngészőben a beszéd felismerés és fordítás indítása után ki kell választani, hogy a beszéd felismerés eredménye, a fordítás eredménye vagy mindkettő jelenjen-e meg.



- Beállíthatóak az előzőleg felismert és lefordított, valamint az aktuálisan felismert és lefordított szövegek stílusaik és méretei:**

-  gomb - félkövér (bold) stílus
-  gomb - dőlt (italic) stílus
-  gomb - aláhúzott stílus
-  gomb - a szöveg betűszínének beállítása
-  gomb - a szöveg háttérszínének beállítása
-  gomb - betűméret

- A beállításokat a jobb oldali bemutató szövegek segítségével lehet ellenőrizni
- A jobb oldali bemutató szövegek háttér felületére jobb egér gombbal kattintva beállítható a HTML megjelenítés háttérszíne



- Az előzőleg felismert és lefordított szövegek megjelenítésének száma**
 - Beállítható, hogy az előzőleg felismert és lefordított szövegekből hány sort szeretnénk megjeleníteni (0 és 10 között)
- Vonalas elválasztás**
 - A vonalas elválasztás bekapcsolása esetén az aktuálisan felismert és lefordított szöveg egy vonallal lesz elválasztva az előzőleg felismert és lefordított szövegtől

5. Szegmentálás

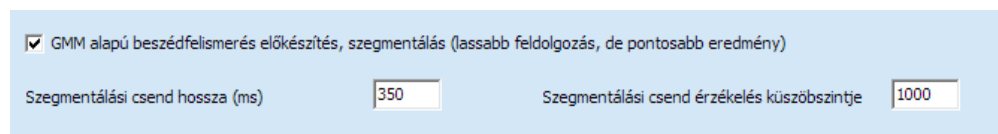
A beszéd felismerő rendszerek számára a beszéd információt szakaszokra kell bontani. Ezen beszéd szakaszokat nem beszéd szakaszok (lélegzet vétel, csend stb.) határolják.

Minél rövidebbek a beszéd szakaszok, annál gyorsabb lesz a szakasz beszéd felismerése (így a felismert szöveg megjelenítése is), de a felismerési pontossága csökkenhet.

Hosszabb beszéd szakaszok esetén lassabb lesz a felismerés, de a felismerés pontossága javulhat.

Egy beszéd szakasz optimális hossza 10-15 másodperc, de minden esetben kevesebb kell legyen, mint 20 másodperc.

Ha egy beszédben nem található 20 másodpercen belül nem beszéd szakasz (lélegzetvételt, csend, stb.), akkor a beszéd 20 másodperces darabokra lesz bontva, még akkor is, ha szó közben kell elvágni a beszédet.



- GMM alapú beszéd felismerés előkészítés, szegmentálás**
 - Bekapcsolásával a beszéd felismerés előkészítése lassabban történik meg, de a beszéd felismerés pontossága javulhat.
- Szegmentálási csend hossza**
 - Itt beállíthatjuk a hangfelvétel beszéd felismerésekor figyelembe vett szünet (csend) hosszát ezredmásodpercben. Az alapértelmezett érték általában megfelelő (350ms). Az alapértelmezett érték módosítását nem ajánljuk.
- Szegmentálási csend érzékelés küszöbszintje**
 - Itt beállíthatjuk azt a jelszintet, amely jelszint alatti jelet csendnek tekint a rendszer és ennek a figyelembe vételével szegmentálja (darabolja) a hangfelvételt. Az alapértelmezett érték általában megfelelő (1000), de szükség esetén próbákkal más érték is beállítható. A beállítható érték a 0 és 65535 között lehet.

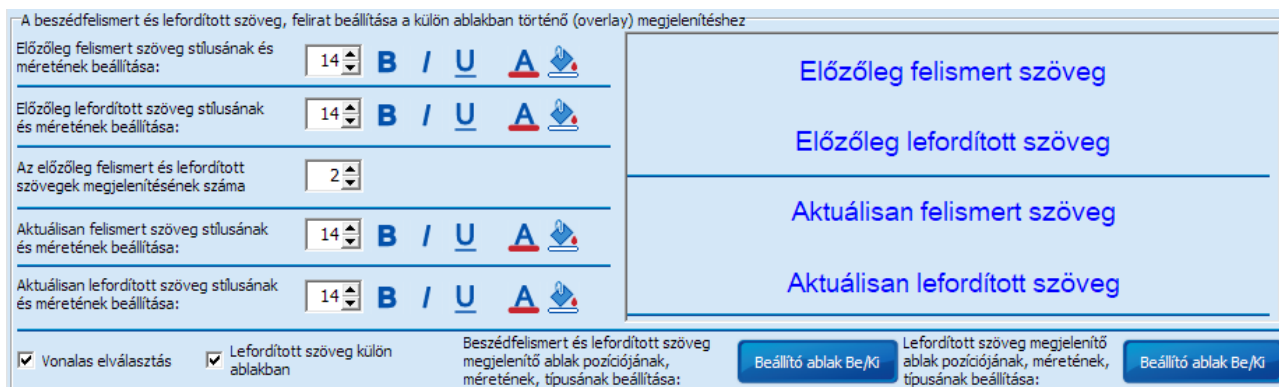
6. Feliratozás megjelenítése

Itt a beszéd felismerés és fordítás eredményének külön ablakban (overlay) történő megjelenítéséhez szükséges beállításokat végezhetjük el.







A beszéd felismerés és fordítás külön ablakban történő megjelenítése esetén a számítógép képernyőjén (desktop) egy vagy két külön ablakban jeleníthetjük meg felismert és lefordított szöveget. A megjelenítő ablakok mindig láthatók, még az ABSONIC felületének háttérbe kerülése esetén is.

Kettő megjelenítő ablak használata esetén az egyikben a beszéd felismert szöveget, a másikban a lefordított szöveget jelenítjük meg.

Az ablakok tetszőleges méretre és pozícióba állíthatók, a háttérük átlátszóvá is tehető, így bármilyen háttérfelületre illetve képre rátehető a szöveg (feliratozás esetén hasznos) vagy egy nézőközönség/hallgatóság látómezőjének felső részébe, akár az előadó fölé is kivetíthető.



- Beállíthatóak az előzőleg felismert és lefordított, valamint az aktuálisan felismert és lefordított szövegek stílusai és méretei:**

-  gomb - félkövér (bold) stílus
-  gomb - dőlt (italic) stílus
-  gomb - aláhúzott stílus
-  gomb - a szöveg betűszínének beállítása
-  gomb - a szöveg háttérszínének beállítása
-  gomb - betűméret

- A beállításokat a jobb oldali bemutató szövegek segítségével lehet ellenőrizni
- A jobb oldali bemutató szövegek háttér felületére jobb egér gombbal kattintva beállítható a megjelenítő ablakok háttérszíne



- Az előzőleg felismert és lefordított szövegek megjelenítésének száma**
 - Beállítható, hogy az előzőleg felismert és lefordított szövegekből 0 és 10 között hány sort szeretnénk megjeleníteni

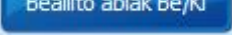
□ **Vonalas elválasztás**

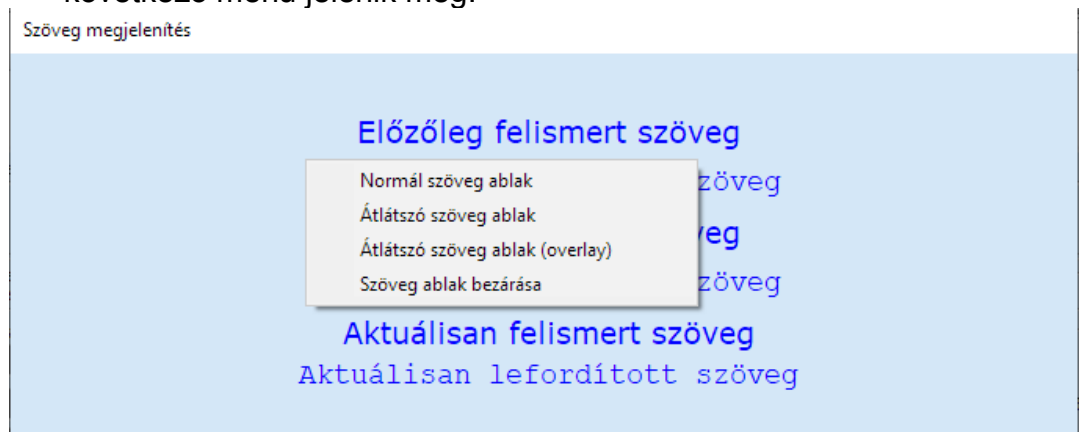
- A vonalas elválasztás bekapcsolása esetén az aktuálisan felismert és lefordított szöveg egy vonallal lesz elválasztva az előzőleg felismert és lefordított szövegtől

□ **Lefordított szöveg külön ablakban**

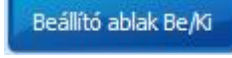
- Bekapcsolása esetén a lefordított szöveg egy másik (overlay) ablakban lesz megjelenítve

□ **Beszéd felismert és lefordított szöveg megjelenítő ablak pozíciójának, méretének, típusának beállítása**

- A  gombra kattintva megjelenik a beállító ablak
- A beállító ablak fejlécén a bal egérgombot lenyomva tartva az ablakot abba a pozícióba lehet mozgatni, ahol a megjelenítéshez szükséges lesz.
- A beállító ablak széleire vagy sarkára kattintva és az egérgombot lenyomva tartva az ablak méretét lehet beállítani.
- A beállító ablak felületére jobb egér gombbal kattintva a következő menü jelenik meg:



- Normál szöveg ablak – kiválasztásával egy normál, a beállított háttérszínnel rendelkező szöveg ablak fog megjelenni
- Átlátszó szöveg ablak – kiválasztásával egy átlátszó szöveg ablak jelenik majd meg, ami az átlátszósága ellenére nem engedi az alatta levő, átlátszó felület használatát, kattintását
- Átlátszó szövegablak (overlay) – kiválasztásával egy olyan átlátszó szöveg ablak jelenik majd meg, ami tényleg csak a szöveget jeleníti meg és az alatta levő felület kattintható - használható marad.
- Szövegablak bezárása – kiválasztásával a beállító ablak bezáródik és méret, pozíció, típus beállítások eltárolódnak

- A  gombra újra rákattintva bezárható a megnyitott beállító ablak, a beállítási adatok tárolásával.

□ **Lefordított szöveg megjelenítő ablak pozíciójának, méretének, típusának beállítása:** akkor használjuk ezt a parancsot, ha a fordítás eredményét másik ablakban szeretnénk megjeleníteni, mint a beszéd felismerés eredményét. Beállítási lehetőségei megegyeznek az előzőkkel.

7. ABSONIC beszédfelismerés (offline) beállításai (ez a szolgáltatás csak egyedi megrendelésre készül el és érhető el azzal, hogy a DEMO verzióban egy kötött szöveganyagon kipróbálható)

Itt az ABSONIC beszédfelismerő rendszer (offline) beállításait végezhetjük el.

ABSONIC beszédfelismerés beállításai

Beszédfelismerés nyelve: Magyar

Beszédfelismerő program neve: C:\Users\Public\ABSONIC2\srn_absonic\bin\ASR.exe

ABSONIC audio port: 5530

ABSONIC modul port: 10500

Pontozás szerinti színezés

Pontozás szerinti színezés felső küszöbhatára: 400

Színezés alsó küszöbszintje: 150

Tallózás

- **Beszédfelismerés nyelve**
 - Itt választjuk ki azt a nyelvet, amelyen a beszédfelismerés történni fog.
- **Beszédfelismerő program neve**
 - Az ABSONIC beszédfelismerő program futtatható állományát kell kitallózni és beállítani.
- **ABSONIC audio port**
 - Az ABSONIC és a beszédfelismerő program közötti audió adatok átvitelét biztosító port száma. Alapértelmezetten: 5530
- **ABSONIC modul port**
 - Az ABSONIC és a beszédfelismerő program közötti vezérlő, és állapot jelző utasítások átvitelét biztosító port száma. Alapértelmezetten: 10500
- **Pontszám szerinti színezés**
 - Bekapcsolása esetén a felismert szöveg a beszédfelismerő rendszer által becsült felismerési pontosság alapján az ABSONIC RTF szövegszerkesztőben színjelöléssel lesz ellátva.
 - A valószínűleg helyesen felismert szöveg betű és háttérszíne változatlan.
 - A bizonytalanul felismert szöveg háttérszíne sárga vagy piros a bizonytalanság szintjétől függően.
 - Ez a funkció csak az ABSONIC offline (?) beszédfelismerőben használható.
- **Pontozás szerinti színezés felső küszöbhatára**
 - Beállítható a pontszám szerinti színezés bekapcsolása esetén az a ponthatár, ami felett a felismert szöveg valószínűleg helyesen felismert szövegnek tekinthető. Ez alatt a felismert szöveg felismerési pontossága bizonytalan. A beállítható értéket ezrelékben, próbák után kell megadni (1-1000).
- **Színezés alsó küszöbszintje**
 - Beállítható a pontszám szerinti színezés alsó küszöbszintje, ami alatt nagyon bizonytalan a felismert szöveg pontossága, felette pedig részben bizonytalan. A beállítható értéket ezrelékben kell megadni (1-1000).

8. ABSONIC beszédfelismerés (online) beállításai

ABSONIC (online) beszédfelismerés beállításai

Beszédfelismerő host név vagy IP cím

ABSONIC audio port ABSONIC modul port

Pontszám szerinti színezés Pontozás szerinti színezés felső küszöbhatára Színezés alsó küszöbszintje

- **Beszédfelismerő host név vagy IP cím**
 - A beszédfelismerő programot futtató számítógép neve vagy IP címe.
- **ABSONIC audio port**
 - Az ABSONIC és a beszédfelismerő program közötti audió adatok átvitelét biztosító port száma. Alapértelmezetten: 5530
- **ABSONIC modul port**
 - Az ABSONIC és a beszédfelismerő program közötti vezérlő és állapot jelző utasítások átvitelét biztosító port száma. Alapértelmezetten: 10500
- **Pontszám szerinti színezés**
 - Bekapcsolása esetén a felismert szöveg a beszédfelismerő rendszer által becsült felismerési pontosság alapján az ABSONIC RTF szövegszerkesztőben színjelöléssel lesz ellátva.
 - A valószínűleg helyesen felismert szöveg betű és háttérszíne változatlan.
 - A bizonytalanul felismert szöveg háttérszíne sárga vagy piros a bizonytalanság szintjétől függően.
 - Ez a funkció csak az ABSONIC (offline?) beszédfelismerésben használható.
- **Pontozás szerinti színezés felső küszöbhatára**
 - Beállítható a pontszám szerinti színezés bekapcsolása esetén az a ponthatár, ami felett a felismert szöveg valószínűleg helyesen felismert szövegnek tekinthető. Ez alatt a felismert szöveg felismerési pontossága bizonytalan. A beállítható értéket ezrelékben kell megadni (1-1000).
- **Színezés alsó küszöbszintje**
 - Beállítható a pontszám szerinti színezés alsó küszöbszintje, ami alatt nagyon bizonytalan a felismert szöveg pontossága, felette pedig részben bizonytalan. A beállítható értéket ezrelékben kell megadni (1-1000).

9. Google beszédfelismerés (online) beállításai

Google beszédfelismerés beállításai


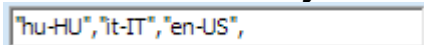
Google API kulcs [Google regisztráció](#)

Beszédfelismerés nyelve > Automatikus nyelvválasztás Beszélő azonosítás

Beszédfelismerés típusa Beszédfelismerési profil Írásjelezés

Felvételi eszköz típusa Mikrofon távolsága Saját szókészlet

- **Google API kulcs**
 - Itt azt a Google API kulcsot kell megadni, ami a Google Cloud Console-ban a beszédfelismerés projekthez létrejött.
 - A **Google regisztráció** gombra kattintva a böngészőben megnyílik az az oldal, ahol a Google API kulcs beszerzéséhez szükséges regisztráció elvégezhető, abban az esetben, ha nem rendelkezik még API kulccsal.
- **Beszédfelismerés nyelve**

-
- Itt választjuk ki azt a nyelvet és nyelvjárást, amelyen a beszédfelismerés történni fog.
 - **Automatikus nyelvválasztás**
 - Bekapcsolása esetén három megadott nyelv közül, automatikusan fogja kiválasztani a beszédfelismerő rendszer azt nyelvet, amelyen a beszédfelismerés történni fog.
 - A  gombra kattintva tudjuk az automatikus nyelvválasztásban használni kíván nyelvet hozzáadni a három használható nyelvhez, a **Beszédfelismerés nyelve** listából kiválasztva.

 - Törölni egyszerűen, a kiválasztott nyelveket tartalmazó szöveg mezőből lehet a nem szükséges nyelvet.
 - **Beszélő azonosítás**
 - A beszélő azonosítás bekapcsolásával a beszédfelismerő rendszer képes - a megfelelő minőségű hangfelvétel esetén - maximum 6 személy közül beazonosítani az aktuális beszélőt. A személyek a felismert szövegben számozással azonosíthatók.
 - A beszélő azonosítás jelenleg csak angol, német, spanyol, francia, olasz, japán és kínai nyelveken működik!**
 - **Beszédfelismerés típusa**
 - Itt a beszédfelismerésre küldött hangfelvétel jellege szerint lehet beállítani a beszédfelismerés típusát, amely lehet:
 - Általános beszélgetés
 - Megbeszélés, találkozó
 - Prezentáció, előadás
 - Telefonbeszélgetés
 - Hangüzenet
 - Rádió, TV műsor
 - Hang keresés
 - Hang vezérlés
 - Diktálás
 - **Beszédfelismerési profil**
 - Itt a beszédfelismerés profilja állítható be, amely lehet:
 - Általános beszédfelismerés
 - Parancsszavak beszédfelismerése
 - Videóról történő beszédfelismerés
 - Telefonbeszélgetés (normál beszédfelismerési minőség)
 - Telefonbeszélgetés (magas beszédfelismerési minőség)
 - **Írásjelezés**
 - Bekapcsolása esetén a beszédfelismerő rendszer – nyelvtől függően korlátozott minőségben – írásjelekkel látja el a felismert szöveget (mondat eleji nagybetű, vessző, felkiáltójel, pont stb.)
 - Az írásjelekkel ellátott beszédfelismert szöveg jelenleg csak angol, német, spanyol, francia, olasz, japán, kínai, orosz és portugál nyelveken működik!**
 - **Felvételi eszköz típusa**
 - A beszédfelismerő rendszerrel közölhetjük a felvételi eszköz típusát illetve a felvételi helyszínt, amely lehet:
 - Ismeretlen

-
- Mobil telefon
 - Számítógépre csatolt mikrofon
 - Vezetékes telefon
 - Járműben
 - Szabadtér, kültér
 - Szoba, beltér
- **Mikrofon távolsága**
 - A beszéd felismerő rendszerrel közölhetjük a felvételhez használt mikrofon távolságát, amely lehet:
 - Ismeretlen
 - Közeli
 - Közepes
 - Távoli
 - **Saját szókészlet**
 - Bekapcsolása esetén a beszéd felismeréshez saját szókészletet használhatunk, amely javíthatja a beszéd felismerés pontosságát. Ez a szókészlet, egy az ABSONIC konfigurációs mappájában található és aSRNSpeechContext.xml nevű állomány szerkesztésével bővíthető.
 - Használata jelenleg nem javasolt**